

## &gt;LGR4 nucleotide sequence (SEQ ID NO:01)

ATGCCGGGCCGCTAGGGCTGCTCTGCCCTGGGCTGCTCGGCCGGGCCAGCGCGCGGCCGCT  
 CTCAGCGCAGCTGGACGGCAGCCGTCGGTGGACTGCTCCGAAAGGGGTTGACGGCGTACCGGAGGGT  
 TTTCTAGAGGAGCTACAACGGCTGGTAACGACCTTCTTATCCATCCAAAAGCCTGTCTGGCTGAAAGAACTCAA  
 GTCCTAACACTCCAGAATAATCAGTTGAGAACAGTGCCCAGTGAAGCCATTACGGACTGAGTGCTTGCAGTCTTACGC  
 TTAGATGCCAACCATATTACCTCAGTCCCAGGAGCAGTTGAAGGGCTTGTCCAGTACGCCATCTGTGGCTGGATGAC  
 AACAGCTTGACGGAAGTGCCCGTGTCCCTCAGCAACCTGCCAGGCGCTGACCTGGCTCTAACAAACATC  
 TCAAGCATTCTGACTTCGCTTACCAACCTTCAAGCTTGTGGTTCTGCATCTGCATAACAATAAAATTAAAGCCTC  
 AGTCAACACTGTTGATGGACTAGATAACCTGGAAACCTTGGACTTGAATTACAATTACTGGATGAGTTCTCAGGCT  
 ATTAAAGCCCTTCCCAGCCTAAAGAGCTGGATTTCACAGTAATTCTATTCTGTATTCTGATGGAGCATTGGTGGT  
 AATCCACTGCTAAGAACTATTCTTGTATGATAATCCTCTGTCTTGTGGGAACCTCAGCATTCAACACCTGTCTGAT  
 CTGCATTGCTTAGTCATTGTTGCAAGCTGGTGCAGTGGTCCCAATCTGACCGAACTGTCCATTGGAGAGTCTA  
 ACCTTGACAGGGACAAAAATAAGCAGCATAACCTGATGATCTGTGCCAAAACAAAAGATGCTGAGGACTCTGGACTTATCT  
 TATAACAATATAAGAGACCTCCAAGTTTAATGGTTGTCGTGCATTGGAAGAAATTCTATTGAGCCTGAAATCTCC  
 CTAATAAAAGAAAATACTTCTAAGGCCAACATCTCTAAGGATTCTAGATCTGAGTAGAAACCTGATCCGTAAATTCA  
 AGTGGAGCTTTCGCAAGCTGGGACAATTACTAACCTGGATGTAAGTTCAATGAATTACTCATTTCTACGGAGGC  
 CTTAAATGGGCTCAATCAACTAAAGCTTGTGGTAACTTCAAGCTGAAAGACGCCCTGGCAGCCAGAGACTTGTCAATCTC  
 AGGTCTCTATCAGTACCATATGCTTATCAGTGTGTGCATTGGGGTGTGACTCTTATGCAAATTAAACACAGAAGAT  
 AACAGCCCCAAGAACACAGTGTGACAAAAGAGAAAGGTGCTACAGATGCAGCAAATGTCACCAGCAGTGTGAAACGAA  
 GAACATAGCCAATAATTATCCACTGTACACCTCAACAGGTGCTTCAAGCCCTGTGAATTACTGGGAGCTGGATG  
 ATTGCGCTTACAGTGTGGTCATTTCCTGGTCCCTGCTTTCAACCTGTTGCTATTAAACAGTGTGGCTTGT  
 TCATCACTGCTGCCCTCAAACCTTCATAGGCTTGATTCTGTCTAACCTACTCATGGGATCTATACTGGCATCCT  
 ACTTTCTTGATGCTGTCTGGCGATTGCGAATTGGCATTGGGGAAACTGGCAGGGCTGCAAGGTAGCC  
 GGGTCTGAGCTTCTCCCTCAGAGAGCGCTGTATTCTATTAAACACTGGCAGTGTGGAAAGAAGCGTATTGCAAAG  
 GATTGATGAAACACGGGAAGAGCAGTCACCTCAGACAGTTCCAGGTGGCCCTCTAGCTTGTGGCTGGCGCAGTG  
 GCAGGCTGCTCCCCCTTTCCACGGAGGGCAATTCTGCATGCCCTTGTGTTGCCGTTCTACAGGAGAAACCCCA  
 TCGTTAGGATTCACTGTGACCTAGTGCTATTAAACTCACTGGCATTAACTATGGCATTACTACACTAAACTATAC  
 TGCAACTTAGAGAAGGAGGACCTGCGAAAACCTCCAGTCTAGCGTGAATTAGCACGTTGCCTGCTCATCTCACAAAC  
 TGCACTTCTCTGCCCTGTTGCATTTCATTCGACCATTGATCACGCAATCTCCATCAGCCCGAGATAATGAAG  
 TCTGTTACACTGATATTCTTCCGTTGCCCTGCTTGCTGAATCCGCTGTATGTTCTCAACCCAAAGTTAAAGAA  
 GACTGGAAGCTACTGAAGCGCGTGTACCAAGGAAACACGGATCTGTTCACTCAGCAGCCAAGGCCGTTGTGG  
 GAACAGGAGTTCTACTATGACTGTGGCATGTATTCCACTTGCAAGGTAACCTGACTGTGACTGCTGTGAGTCATT  
 CTTTGACAAAACAGTATCATGCAACACTTAATAAAATCGCACAGTTGCTCTGTATTGACAGCGGCCTTGTCCAGAGG  
 CCAGAGGCCTACTGGTCTGATTGTGGTACACAGTCAGCCCATTCTGACTATGCAAGATGAAGAAGATTCTTGTCTCAGAC  
 AGCTCTGACCAAGGTGCAGGCCGTGGACGAGCCTGCTTCTACAGAGTCGTGGATTCCCTCTGGTGCCTATGCTTACAAT  
 CTACAGAGAGTCAGAGACTGA

## &gt;LGR4 amino acid sequence (SEQ ID NO:02)

MPGPLGLCFLALGLLSAGPSGAAPPLCAPCSDGDRRVDGKGLTAVPEGLSAFTQALDISMNNITQLPEDAFKSF  
 FLEELQLAGNDLSLIHPKALSGLKELKVLTQNNQLRTVPSEAIHGLSALQSLRDLANHITSVPEDSFEGLVQLRHLWLD  
 NSLTEVPVRPLSNLPTLQALTALNNISSIPDFAFTNLSSLVLHNNKIKSLSQHCFDGLDNLETLDLNLYNLD  
 FQPLQKSLRILDLRNLI  
 REIH  
 SGAFAKLGTITNLDSFNELETSFPTEGNLNLQQLKLVGNFKLKDALAARDFANLRSLSVPYAYQCCAFWGCDSLCKLN  
 TED  
 NSPQEHSVTKEKGATDAANVTSTAENEHSQII  
 IHCTPSTGAFKPCEYLLGSMWIRLTWVFIFLVALLFNLLV  
 ILTVFASC  
 SSLPASKLFIGLISVSNLLMGIYTGILTFLDAVSWGRFAEFGIWETGSGCKVAGSLAVFSSES  
 AVFLLTAAVERS  
 VFAK  
 DLMKHGKSSHRLRQFQVAALLALLGA  
 AAVAGCFPLFHGGQYSASPLCLP  
 FPTGETPSLGF  
 T  
 V  
 L  
 V  
 L  
 L  
 N  
 S  
 L  
 A  
 F  
 L  
 L  
 M  
 A  
 I  
 I  
 Y  
 T  
 K  
 L  
 Y  
 CNLEKEDLSENSQSSVI  
 KHVAWL  
 I  
 F  
 T  
 N  
 C  
 I  
 F  
 F  
 C  
 P  
 V  
 A  
 F  
 F  
 S  
 A  
 P  
 L  
 I  
 T  
 A  
 I  
 S  
 I  
 S  
 P  
 E  
 I  
 M  
 K  
 S  
 V  
 T  
 L  
 I  
 F  
 F  
 P  
 L  
 P  
 A  
 C  
 L  
 N  
 P  
 V  
 L  
 Y  
 V  
 F  
 F  
 N  
 P  
 K  
 F  
 K  
 E  
 DWKLLKRRVTRKHGSVS  
 VSISSQGGCGE  
 QDFYYDCGMYSHLQGNLT  
 VCDCCES  
 FLLTKPVSKH  
 LIKSHSCP  
 VLTAASCQR  
 PEAYWSDCGTQS  
 AHDYADEEDSF  
 VSDSSD  
 QVACGRAC  
 FYQS  
 RGF  
 PLV  
 RYAYN  
 LQR  
 VRD

FIG. 1

>Nucleotide sequence of LGR5 (total 2082 nucleotides) (SEQ ID NO:03)

CTACATCTCCATAACAATAGAACATCCACTCCCTGGGAAAGAACATGCTTGATGGGCTCCACAGCCTAGAGACTTAGATTTA  
 ATTACAATAACCTGATGAATTCCCCACTGCAATTAGGACACTCTCCAACCTAAAGGAACATAGGATTCTATAGCAACAAAT  
 ATCAGGTCGATACTGAGAAAGCATTGAGGCAACCCCTCTTATTACAATACATTCTATGACAATCCATCCAATT  
 GTGGGAGACTGCTTTCAACATTACCTGAACAAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAATTCCCT  
 GATTTAACTGGAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATCTCATCTCTCCTCAAACCGTCTGCAAT  
 CAGTTACCTAATCTCCAAGTGCTAGATCTGCTTACAACCTATTAGAAGATTACCCAGTTTCAGTCTGCCAAAGCTT  
 CAGAAAATTGACCTAACAGACATAATGAAATCTACGAAATTAAAGTTGACACTTCCAGCAGTTGCTTAGCCTCCGATCGCTG  
 ATTGGCTTGAACAAAATTGCTATTATTACCCCCAATGCATTTCACCTTGCCATCCCTAATAAGCTGGACCTATCG  
 TCCAACCTCCTGCTGCTTTCTATAACTGGTTACATGGTTAACTCACTTAAATTACAGGAATCATGCTTACAG  
 AGCTGGATATCATCTGAAAACCTTCCAGAACACTCAAGGTXATAGAAATGCCTTATGCTTACAGTGCTGTGCAATTGGAGTG  
 TGTGAGAATGCCATAAGATTCTAATCAATGGAATAAAGGTGACAACAGCAGTATGGACGACCTTCATAAGAAAGATGCT  
 GGAATGTTCAAGGCTAACAGATGAACGTGACCTGAAGATTCTGACTTTGAGGAAGACCTGAAAGCCCTCATTCA  
 GTGCAGTGTTACCTTCCCCAGGCCCCCTCAAACCTGCTGAACACCTGCTTGATGGCTGGCTGATCAGAATTGGAGTGTGG  
 ACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTGGTGAACCTCAACAGTTTCAGATCCCTCTGTCATATTCCCCCATT  
 AAACGTAAATTGGGTCTCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGCCGTGCTGGCTGGATGCGTTC  
 ACTTTGGCAGCTTGCACGACATGGTGCCTGGTGGGAGAATGGGGTTGGTGCCTATGTCATTGGTTTTGTCCATT  
 GCTTCAGAACATCTGTTTCTGTTACTCTGGCAGCCCTGGAGCGTGGGTTCTCTGTGAAATATTCTGCAAATTGAA  
 ACGAAAGCTCCATTCTAGCCTGAAAGTAATCATTTCCTGTCCTGCCCCCTGCTGGCCTTGACCATGGCCAGTCCCCCTG  
 CTGGGTGGCAGCAAGTATGGCCCTCCCTCTGCCTGCCTTGCCTTTGGGAGCCCAGCACCATGGCTACATGGTC  
 GCTCTCATCTGCTCAATTCCCTTGCTTCCATGATGACCATGGCTTACACCAAGCTCTACTGCAATTGGACAAAGGG  
 GACTCTGGAGAATATTGGGACTGCTCTATGGAAAACACATTGCCCCTGCTTCCACCAACTGCACTCTAAACTGCCCT  
 GTGGCTTCTGCTCTCCTCTTAAACACCTTACATTATCAGTCTGAAAGTAATTAAAGTTATCCTCTGGTGGTA  
 GTCCCACCTCCTGCATGCTCAATTCCCTCTACATTTGTTCAATCCTCACTTTAAGGAGATCTGGTGGACCTGAGA  
 AAGCAAACCTACGTGAGAACAGATCAAACACCCAAAGCTTGATGTCAATTAACTCTGATGATGTCGAAAACAGTCCTG  
 GACTCAACTCAAGCCTGGTAAACCTTACAGCTCCAGCATCACTTATGACCTGCTCCAGTCCGTGCCATACCAAGCT  
 TATCCAGTGACTGAGAGCTGCCATTTCTGTCATTGTCCCAGTCTCTAA

>amino acid sequence of LGR5 (total 693 amino acids) (SEQ ID NO:04)

LHLHNNRIHSLGKKCFDGLHSLETLDLNYYNNLDEFPTAIRTLSNLKELGFHSNNIRSIPEKAVGNPSLITIHFYDNPIQF  
 VGRSAFQHLPFLRTLNGASQITEFPDLTGATNLESLLTGAQISSLPQTVNCNQLPNLQVLDLSYNLLEDLPSFSVCQKL  
 QKIDLRHNEIYEIKVDTFQQLLSRLSLNLAWNKIAIIPNAFSTLPSLIKLDLSSNLLSSFPITGLHGLTHLKLGNHALQ  
 SWISSENFPELKVIEMPTYAYQCCAFGVCEAYKISNQWNKGDNSMDDLHKKDAGMFQAQDERDLEDFLLDFEEDLKALHS  
 VQCPSPGPFKPCEHLDGWLIRIGVWTIAVLALTCNALVTSTVFRSPLYISPPIKLLIGVIAAVNMLTVSSAVLAGVDAF  
 TFGSFARHGAWENGVGCHVIGFLSIFASESSVFLTLAALERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPL  
 LGGSKYGASPLCLPLPFGEPSTMGYMVALILLNSLCFLMMTIAYTKLYCNLDKGDLENIWDCSMVKHIALLFTNCILNCP  
 VAFLSFSSLINLTFISPEVIKFILLVVVPLPACLNPLLYILFNPHFKEDLVSRLRKQTYWTRSKHPSLMSINSDDVEKQSC  
 DSTQALVTFTSSSITYDLPPSSVPSPAYPVTECHLSSAFVPC

FIG. 2

>Final LGR7 (LGR7-Long variant) full length sequence (2467 nt) (SEQ ID NO:05).

GAAAGGAGGAAAGAAAAAAAGAGGAATGAAAGAGACAGAGAAAGGAATGGGACTGGAAGGAGGGAGGACTGCTTT  
 GTAACTGCTAAGATTGCAAGACAGAAATAGCACACAAACACTGTGAGCTGTGATCGATCAGAAACCAAGACCAAATT  
 TTGCTCACTTCATTAATCAGTTGCTCAGATAGAAGGAATGACATCTGGTTCTGTCTTCTACATCTTAATT  
 TGGAAAATATTTTCTCATGGGGTGGACAGGATGCAAGTGTGCTCCCTGGCTATTCCCTGTGGAACATCACAA  
 AGTGTGCTGCCTCAGCTCCTGCACTGTAACGGTGTGGACGACTGCGGGAACTCAGGCCATGAGGACAACGTGGAGAC  
 ACAATGGATGGTCCATGCAATTGACAATATTGCCCAGTTACTACAAAATGACTTCCAATATCCTTTGAGGC  
 AGAAACACTGAATGTTGGTGGTCTGTGCCAGTGCATGCAATGTCCTTGCCAAGGTCTGGAGCTTGACTGTGATGAAA  
 CCAATTACGAGCTGTTCCATCGGTTCTCAAATGTGACTGCAATGTCATTGAGACTTAAAGATTACATCCATCTCCAT  
 CCTCCTGATTGCTTCAGAAATTATCATGATCTCAGAACGACTGTACCTGCAAAACAATAAGATTACATCCATCTCCAT  
 CTATGCTTCAGAGGACTGAATAGCCTTACTAAACTGTATCTCAGTCATAACAGAATAACCTCTGAAGCCGGTG  
 TTTTGAAAGATCTTACAGACTAGAATGGCTGATAATTGAAGATAATCACCTCAGTCGAATTCCACCAACATT  
 TATGGACTAAATTCTCTTATTCTTAGTCTGATGAATAACGTCCTCACCGTTACCTGATAAACCTCTGTCA  
 ACACATGCAAGACTACATTGGCTGGACCTTGAAGGCAACCATACTCCATAATTAAAGAAATTGACTTTATTCC  
 GCAGTAATTAACTGTTTAGTGTGAGGAAAACAAAATTAACTTAAATGAAAATACTTTGACCTCTCCAG  
 AAACGGATGAATTGATTAGAAGTAATAAGATTGAAAATCTTCACCGCTTATATTCAAGGACCTGAAGGAGCT  
 GTCACAATTGAATCTTCTATAATCCAATCCAGAAAATTCAAGCAAACCAATTGATTATCTTGCAAACTCAAGT  
 CTCTCAGCTAGAAGGGATTGAAATTCAAATATCCAACAAAGGATGTTAGACCTTATGAAATCTCTCACATA  
 TATTAAAGAAATTCACTGACTGTGGTATGCACCATGTTCGCAGCTGTAAACCAAACACTGATGGAATTCACT  
 TCTAGAGAATCTTGGCAAGCATTTCAGAGAGTATTGTCGGTTGTATCTCAGTTACCTGCTTGGAAACA  
 TTTTGTCATTGCACTGCGACCTTATATCAGGTCTGAGAACAGCTGTATGCCATGCAATCATTCTCTGCTGT  
 GCCGACTGCTTAATGGAATATATTATTCTGATCGGAGGCTTGACCTAAAGTTCTGGAGAACATAAAAGCA  
 TGCGCAGCTGTGGATGGAGAGTACTCATTGTCAGCTGTAGGATCTTGGCCATTCTGTCACAGAAGTATCAGTT  
 TACTGTTAACATTCTGACATTGGAAAATACATCTGATTGCTATCCTTTAGATGTCAGGACCTGGAAAATGC  
 AGAACAAATTACAGTTCTGATTCTCATTGGATTACTGGTTTATAGTGGCTTCTCATTGAGCAATAAGGAATT  
 TTTCAAAAACACTATGGCACCAATGGAGTATGCTTCCCTCTTCATTGAGAACAGAACAGTATTGGAGGCCAGA  
 TTTATTCACTGGCAATTCTGGATTAAATTGGCCGATTATCATAGTTTCTATGGAAGCATGTT  
 TATAGTGTTCATCAAAGTGCATAACAGCAACTGAAATACGGAAATCAAGTTAAAAGAGATGATCCTGCCAACG  
 TTTTTCTTATAGTATTACTGATGCATTGCTGGATACCCATTGGTAGTGAATTCTTCACTGCTTCAGG  
 TAGAAATACCAGGTACCATACCTCTGGTAGTGAATTCTGCCCATTACAGTGTGTTGAACCCAATTCTC  
 TATACTCTGACCACAAGACCATTAAGAAATGATTATCGGTTTGGTATAACTACAGAACAAAGAAAATCTATGGA  
 CAGCAAAGTCAGAAAACATATGCTCCATATTGATCTGGTAGAATGTGGCCACTGCAGGAGATGCCACCTGAGT  
 TAATGAAGCCGGACCTTTCACATACCCCTGTGAAATGTCACTGATTCTCAATCAACGAGACTCAATTCTATTCA  
 TGA

>Final LGR7 (LGR7-long variant, total 757 amino acids)(SEQ ID NO:06)

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQLLHCNGVDDCGNQADEDNCGDNNNGWSMQFDKYFA  
 SYYKMTSQYPFEAETPECLVGSVPVQCLCQGLELDCDETNLRAPSVSSNVNTAMSLQWNLIRKLPDCFKNYHDLQK  
 LYLNQNNKITSISIYAFRGLNSLTKLYLSHNRITFLKPGVFEDLHRLEWLIEEDNHLRSISPPTFYGLNSLILLVLMN  
 NVLTRLPDPKPLCQHMPRLHWLDLEGNHIHNLRLNTFISCSNLTVLVMRKKNINHLENFTAPLQKLDELDLGSNKIE  
 NLPPLIFKDLKELSQNLNSYNPIQKIQANQFDYLVKLKSLSLLEGIEISNIQORMFRPLMNLSHIYFKKFQYCGYAPH  
 VRSCKPNTDGISLLENLLASIIQRFVWWVSAVTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIVLFVIG  
 GFDLKFRGEYNKHAQLWMESTHCQLVGSLAILSTEVSVLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILLIWITG  
 FIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIYSAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEI  
 RNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTTTSWVVFILPINSALNPILYTLTTRPFKEMIH  
 RFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPLEMKPDLFTYPCEMSLISQSTRLNSYS\*

## &gt;Final LGR7 (LGR7-Short variant) full length sequence (3584 nt)(SEQ ID NO:07)

CTGCTTGTAACTGCTAAGATTGAGACAGAAAATAGCACACAACCCTGTGAGCTGTGCGATTGAGAAACCAAGA  
 CCAAATTGCTCAGTTCAATTAACTCAGTGCTCAGATAGAAGGAAATGACATCTGGTTCTGTCTTCTACATCT  
 TAATTTGGAAAATATTTCTCATGGGGTGGACAGGATGTCAGTGCTCCCTGGCTATTTCCCTGTGGAAAC  
 ATCACAAAGTGCTTGCTCAGCTCCTGCACTGTAACGGTGTGGACACTGCGGAATCAGGCCGATGAGGACAAC  
 TGTGGTGGTTGTGCCAGTGCACTGCTTGCAAGGCTGGACTGGATGAAACCATTTACGAGTGTCCAT  
 CGGTTCTCAAATGTGACTGCAATGTCAGTGGAACTTAAAGAAAGCTTCCTCTGATTGCTTCAGAAGAAT  
 TATCATGATCTCAGAAGCTGGACCTGCAAAACAATAAGATTACATCCATCTCATCTGCTTCAGAGGACTGAA  
 TAGCCTTAACAAACTGATCTCAGTCATAACAGAATAACCTTCTGAAGCGGGTGTGTTGAAGATCTTCACAGAC  
 TAGAATGGCTGATAATTGAAGATAATCACCTCAGTCGAATTCCCACCAACATTATGGACTAAATTCTTATT  
 CTCTTAGTCTGATGAATAACGTCCTCACCCGTTACCTGATAAACCTCTGTCACACATGCCAAGACTACATTG  
 GCTGGACCTTGAAGGCAACCATATCCATAATTAAGAAATTGACTTTATTCCTGAGTAATTAACTGTTTAG  
 TGATGAGGAAAACAAAATTAACTCAATTAAAGAAATTACTTTGACCTCTCCAGAAACTGGATGAAATTGATT  
 GGAAGTAATAAGATTGAAAATCTCCACCGCTTATATTCAAGGACCTGAAAGGAGCTGTCACAATTGAATCTTCTA  
 TAATCCAATCCAGAAAATTCAAGCAAACCAATTGATTATCTGTCAAACTCAAGTCTCAGCCTAGAAGGGATTG  
 AAATTCAAAATATCAAACAAAGATGTTAGACCTTATGAATCTCTCACATATATTTAAGAAATTCCAGTAC  
 TGTGGGTATGCACCATGTTCGCAGCTGAAACCAACACTGATGAAATTCTCATCTAGAGAATCTTGGCAAG  
 CATTATTCAAGAGAGTATTGCTGGTTGTATCTGAGTTACCTGCTTGAAACATTGTCATTGCTGTGGCAACTGTTA  
 CTTATATCAAGGCTGAGAACAGCTGATGCCATGTCATCATTCTCTGCTGTGGCAACTGTTAATGGAATA  
 TATTATTGCTGATCGGAGGTTGACCTAAAGTTGCTGGAGAAATACAATAAGCATGCGCAGCTGTTAGGAGAG  
 TACTCATTGCTGAGCTGTTAGGATCTTGGCATTCTGTCACAGAAGTATGTTACTGTTAACATTCTGACAT  
 TGGAAAATACATCTGCTATCTTCTAGATGTTGAGACCTGAAACATTGAGAATTTCAAAACTACTATGGCAC  
 CTCATTGATTACTGGTTTATAGTGGCTTCATTGAGAATAGGAATTTCAAAACTACTATGGCAC  
 CAATGGAGTATGCTTCCCTCTTCATTGAGAATAGGAACAGTATTGGAGGCCAGATTTCAGTGGCAATTTC  
 TTGGTATTAATTGGCGCATTATCATCATAGTTCTTCTATGGAGCATGTTATAGTGTCTCATCAAAGTGC  
 ATAACAGCAACTGAAATACGGAATCAAGTAAAAAGAGATGATCCTGCAAACGTTTCTTCTTATAGTATTAC  
 TGATGCATTATGCTGGATAACCATTGTTAGTGAATTCTTCACTGCTTCAGGTAGAAATACCAGGTACCA  
 CCTCTGGTAGTGTATTCTGCCATTAAACAGTGTGTTGAACCCATTCTCTATACTCTGACCACAAGACCA  
 TTTAAAGAAATGATTGATCGGTTGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTGAGAAA  
 TGCTCCATCATTGATGGGCGAAATGTTGCACTGAGGAGATGCCACCTGAGTTATGAAGCCGACCTTCA  
 CATAACCCGTGAAATGTCAGTATTCTCAATCAACGAGACTCAATTCTATTGACTGAAATTGCAACT  
 TCTTCGAGAGAATACTGTTGGGTTGCTTCATGAGGGATTACTGGTATGAAATACCACAAAATTATTT  
 AATAATAGCTAAGATAAAATATTACAAGGACATGAGGAAAATAAAATGACTAATGCTTACAAAGGAAAGTAA  
 TTATATCAATAATGTATATATTAGTAGACATTGCTCATAGAAATTAGAGAAATCTACTTCAGTAACATT  
 CATTCTCTAACATGCAATTGAGTACCCACTACTATGTCATAGCATGCAATATAGTCTGGAAAGTAGACAGT  
 GCAGAACCTTCAACTGTAGATAGTGTAAATGACAAAGACTATACAAGTCATCTGAGTTCTAGTTAAAG  
 TAGAGCTTACCTGTCATGTCATGCAAGAACATAGGCACCTTAAATAAGGTTAAAGTTGGAAACTCA  
 GTGTATTGTCATGAGAACATGAGGAAATGTCAGTGTGTTGCAAAATAATTCTGTTAAGAATCCATCT  
 AAGTTCCATACACTGAGAGCCAACACACATTACTAAAGATGCTTGCTAGAAACTCAAAACAGCA  
 CTTCTTGGCACTTCTGCCAGTTCTTGTAAATGAACATCATCATATGAAATTGAAATAGGAGAGTA  
 TGAGTACGGCAGAGAAGTGGATCAGAAAACACTAGAATGAGGATAAAACATTACATTAGTGGAAACTCT  
 CCCTGTATTGTCAGTTAACTGATTTCAACAAGGATGCCAAGACAAAAGGCTTCAACAAACCGTGTGTT  
 AGAACAGACCTAAGTGGTTAACCTCACCCACTTTAGATGGGTGAATGTTATGGTGTGAAATATCTCAG  
 GTTAAAGAAAAGAGCTGGAATGCACTGATTGAGAACCTTAATTGAGGAAAGGTCTGTATGTACACATT  
 CACTTTAACGAGAAAATCTTCTCAAGAAATGACTTTACTTTCTCTTGCACTGCCAGCACGTGAGATA  
 TTAAACTAGTTGTTCTCTAGTCTACGTTATTAGNATTGCTTCATAATGTGAAACCTTAAAGCAGGAG  
 AAGAAAATGTTTCAGATAGTTCAAATACNCCAAAATGTTGCAACACAAAATCTGAACTNAACCATAAT  
 CCTTATTGAATATAGTTGATAGNTTGTCTGAAAACCC

## &gt;Final LGR7-S ORF (722 amino acids) (SEQ ID NO:08)

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYPFCGNITKCLPQLLHCNGVDDCGNQADEDNCVVLQCMSPGLEL  
 DWMKPFTSVPSVSSNVTAMSLQWNLIRKLPDFKNYHDLQKLDLQNNKITSISIYAFRGLNSLTKLYLSHNRITFL  
 KPGVFEDLHRLEWLIEDNHLSRISPPFTYGLNSLILVLMMNVLTRLPDPLCQHMPLRHLDLEGNHINLRNLT  
 FISCSNLTVLVMRKKNKINHLENFTAPLQKLDDELGSNKIENLPLIFKDLKELSQNLNSYNPIQKIQANQFDYLV  
 KLKSLSLEGIEISNIQQRMFRLPMNLSHYFKKFQYCGYAPHVRSCPKNTDGISLLENLLASIIQRVFWVVS  
 FGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHQLVGS  
 LAILSTE  
 VSVLTTFLTLEKYICIVYPFRCVRPGKCRTITVLILI  
 WI  
 TGTGIVAFIPLSNKEFFKNYYGTNGVCFPLHSED  
 TESI  
 GAQIYSAIFGINLAAFIIVFSYGS  
 MFSYVHQSAITATEIRNQVK  
 KEMILAKRFFFIVFTDALCW  
 WIPIFVV  
 KFLS  
 LLQVEIPG  
 TITSWV  
 VIFILP  
 INSALN  
 PLYLT  
 TTRPF  
 KEMI  
 HRFW  
 YNR  
 QRK  
 SMD  
 SKG  
 QK  
 T  
 YAPS  
 F  
 I  
 W  
 V  
 E  
 M  
 W  
 P  
 L  
 Q  
 E  
 M  
 P  
 D  
 L  
 F  
 T  
 Y  
 P  
 C  
 E  
 M  
 S  
 L  
 I  
 S  
 Q  
 S  
 T  
 R  
 L  
 N  
 S  
 Y  
 \*

5/8

&gt;Alignment of LGR7-L with LGR7-S

**Query=LGR7-L****Sbjct=LGR7-S**

Query: 1 MTSGSVFFYILIFGKYFSHGGQDVKCSLGYFFPCGNITKCLPQLLHCNGVDDCGNQADED 60  
 MTSGSVFFYILIFGKYFSHGGQDVKCSLGYFFPCGNITKCLPQLLHCNGVDDCGNQADED

Sbjct: 1 MTSGSVFFYILIFGKYFSHGGQDVKCSLGYFFPCGNITKCLPQLLHCNGVDDCGNQADED 60

Query: 61 NCGDNNGWSMQFDKYFASYYKMTSQYPFEAETPECLVGSPVQCLCQ---GLELDCDETN 117  
 NC V V C C GLELD +  
 Sbjct: 61 NC-----VVVLCQCMSLPGLELDWMKP- 82

Query: 118 LRAVPSVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKITSISIYAFRGLNSLT 177  
 +VPSVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKL LQNNKITSISIYAFRGLNSLT

Sbjct: 83 FTSPSPVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLQNNKITSISIYAFRGLNSLT 142

Query: 178 KLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRL 237  
 KLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRL

Sbjct: 143 KLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRL 202

Query: 238 PDKPLCQHMPRLHWLDLEGNHINLRLNLTFISCSNLTVLVMRKNKINHLENENTFAPLQKL 297  
 PDKPLCQHMPRLHWLDLEGNHINLRLNLTFISCSNLTVLVMRKNKINHLENENTFAPLQKL

Sbjct: 203 PDKPLCQHMPRLHWLDLEGNHINLRLNLTFISCSNLTVLVMRKNKINHLENENTFAPLQKL 262

Query: 298 DEELDLGSNKIENLPPLIFKDLKELSQLNLNSYNPIQKIQANQFDYLVKLKSLSLLEGIEISN 357  
 DEELDLGSNKIENLPPLIFKDLKELSQLNLNSYNPIQKIQANQFDYLVKLKSLSLLEGIEISN

Sbjct: 263 DEELDLGSNKIENLPPLIFKDLKELSQLNLNSYNPIQKIQANQFDYLVKLKSLSLLEGIEISN 322

Query: 358 IQQRMFRPLMNLSHIYFKKFQYCGYAPHVRSCPNTDGISSLENLLASIIQRVFVVVVA 417  
 IQQRMFRPLMNLSHIYFKKFQYCGYAPHVRSCPNTDGISSLENLLASIIQRVFVVVVA

Sbjct: 323 IQQRMFRPLMNLSHIYFKKFQYCGYAPHVRSCPNTDGISSLENLLASIIQRVFVVVVA 382

Query: 418 VTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIFYLFIGGFDLKFRGEYNKHAQ 477  
 VTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIFYLFIGGFDLKFRGEYNKHAQ

Sbjct: 383 VTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIFYLFIGGFDLKFRGEYNKHAQ 442

Query: 478 LWMESTHCQLVGLAILSTEVSVLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWI 537  
 LWMESTHCQLVGLAILSTEVSVLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWI

Sbjct: 443 LWMESTHCQLVGLAILSTEVSVLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWI 502

Query: 538 TGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVF 597  
 TGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVF

Sbjct: 503 TGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVF 562

Query: 598 SYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEI 657  
 SYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEI

Sbjct: 563 SYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEI 622

Query: 658 PGTITSWVVFILPINSALNPILYTLTTRPFKEMIHRFWNYRQRKSMDSKGQKTYAPSF 717  
 PGTITSWVVFILPINSALNPILYTLTTRPFKEMIHRFWNYRQRKSMDSKGQKTYAPSF

Sbjct: 623 PGTITSWVVFILPINSALNPILYTLTTRPFKEMIHRFWNYRQRKSMDSKGQKTYAPSF 682

Query: 718 IWVEMWPLQEMPPPELMKPDPLFTYPCEMSLISQSTRLNSYS 757  
 IWVEMWPLQEMPPPELMKPDPLFTYPCEMSLISQSTRLNSYS

Sbjct: 683 IWVEMWPLQEMPPPELMKPDPLFTYPCEMSLISQSTRLNSYS 722

6/8  
**FIG. 6****Signal peptide**

LGR4	MPGPLGLLCFLALGLLGSAGPSGA
LGR5	MDTSRLGVLLSLPVLLQLATG
LHR	MKQRFSALQQLKLLLLLQPPLPRA
FSHR	MALLLVSLLAFLSLGSG
TSHR	MRPADLQLVLLDLPRDLGG

**N-flank cysteine-rich sequence**

LGR4	APPL AA-P S DGDR----RVD SGKGLTAVPEGLSAFTQAA
LGR5	GSSPRSGVLLRG P-TH H EPDGRMILLRVD SDLGLSELPNSNLSVFTSY
LHR	LREAL P-EP N VPDG--ALR-- PGPTAGLTR
FSHR	HHRI H SNRVFL----- QESKVTEIPSDLPRNAIE
TSHR	MG SSPP E HQEED--FRVT KDIQRIPSLPPSTQT

**Leucine-rich repeats**

LGR4	DISMNNITQLPED KSFPFLEELQLAGN -- SL HPKALSG KE KVTLTQ -- Q
LGR5	DLSMNNNISQLLPNPLPSLHFLEELRLAGNA-- TY PKGA TG YS KVLMQ -- Q
LHR	SLAYLPVKVIPSQ RGLNEVIKIEISQI S- ER EANA DN LN SEILIQ TK -
FSHR	RFVLTKLRLVIQKG SGFGDLEKIEISQN V- EV EADV SN PK HEIRIEKAN -
TSHR	KLIETHLRTIPSH SNLPNISRIYVSI- VT QQLESHS YN SKVTHIEIR TR -
LGR4	RTV- SE IHG SA QS RLDA H- TSV EDS--FEGLVQLRH WLD S-L- EV VR
LGR5	RHV- TE LQN RS QS RLDA H- SYV P-SC-FSGLHSLRH WLD A-L- E VQ
LHR	RYIE -G FIN PG KY SIC- TG RKF DVTKVFSSESNFI- EIC LHI- T GN
FSHR	LYIN -E FQN PN QY LIS- TG KHL DVHK-IHSLQKVL- DIQ INIH - ERN
TSHR	TYID -D LKE PL KF GIF- TGLKMF DLTQ-VYSTDIFFI EIT PYM- S VN
LGR4	PLSN P-TLQA T AL NISSIPDF T LSS VV H HN K-IKSLSQHC D LDN-LE
LGR5	A RS S-ALQAMT AL KIHHIPDY G LSSWVV H HN R-IHSLGKKC D LHS-LE
LHR	A QGMNNESVT K YG GFEEVQSH - GTT TS E KE VHLEKMHINGA R A-TGPK
FSHR	S VG SFESVI W NK GIQEIHNC - GTQ DE N SD NNLEELPNDV H A-SGPV
TSHR	A QG CNETLT K YN GFTSVQGY - GTK DAVY NK KYLTVIDKDA G VYSGPS
LGR4	T LNYYNYLDEF Q-AIKA PS KELGFHSNSISVI D-GA GGNPL RTIH - DNPLS
LGR5	T LNYYNNLDEF T-AIRT SN KELGFHSNNIRSI E-KA VGNPS ITIHF- DNPIQ
LHR	T ISSTKLQAL SYGLESIQR I-ATS-SYSLKKL SRET V-N-- LEAT T -----
FSHR	I ISRTRIHSN SYGLEN KK R-ARSTYN-LKKL TLEKLVA--- MEAS T -----
TSHR	L VSQTSVTAL SKGLEH KE I-ARNTWT-LKKL LSLS LH--- TRAD S -----
LGR4	FVGNSAFHNLSDLHCLVIRGASLVQWFPNLTGTVHLESLTGTGKISSIPDDLCQNQKML
LGR5	FVGRSAFQHLPELRTLTLNGASQITEFPDLTGTANLESLTGTGAQISSLPPQTCVNQLPNL
LHR	-----
FSHR	-----
TSHR	-----

7/8

LGR4	RTLDLSYNNIRDLPNFNGCRALEEISLQRNQISLIKENTFQGLTSRLILDLSRNLI
LGR5	QVLDLSYNLLEDLPFSVCQKLQKIDLRHNEIYEIKVDTFQQLLSRLS
LHR	-----
FSHR	-----
TSHR	-----

LGR4	SGAFAKLGTITNLDSFNELTSPTEGLNGLNQLK
LGR5	PNAFSTLPSLIKLDLSSNLLSSFPITGLHGLTHLK
LHR	-----
FSHR	-----
TSHR	-----

**C-flank cysteine-rich sequence**

LGR4	LVGNFKLKDALAARDFANLRSLSV	YAYQ	WGCDSDLCKLNTE	DNSPQEHSVTKEKG
LGR5	LTGNHALQSLISSENFP	ELKVIEM	YAYQ	GVCENAYKISNQWNKGDNSSMDDLHK
LHR	-----	-----	--SH	RNLPTKEQNFHSIS
FSHR	-----	-----	--SH	ENFSKQC
TSHR	-----	-----	--SH	ESTVR

LGR4	TDAANVTSTAENE	HS-----	-----	-----
LGR5	DAGMFQAQDERDL	DF-----	-----	-----
LHR	KVSNKTLYSSMLA	SE-----	-----	-----
FSHR	QTRGQRSSLAEDN	SS-----	-----	-----
TSHR	SVNALNSPLHQEY	ENLGDSIVGYKEKSKFQDTHNNAHYYVFFEEQ	EDEIIGFGQELKNP	-----

LGR4	-----	QIIH	T	STGA	K	YLLGSWMI
LGR5	-----	LLDFEEDLKALHSVQ	S	SPGP	K	HLLDGWL
LHR	-----	LSGWDYEGFCLPKTPR	-A	EPDA	N	DIMGYDFL
FSHR	YSRGFDMTYTFDYLDCNEVVDVT	S	KPDA	N	DIMGYNIL	
TSHR	QEETLQAFDSHYDYTICGDSEDMV	T	KSDE	N	DIMGYKFL	

**Transmembrane**

	TM 1	TM 2					
LGR4	LTV F	FLV LLF LL	ILTVFA	CSS PASKLFIGLISVSNLLM	IYTGILTFL	AVSW	
LGR5	IGV T	AV LTC AL	TSTVFR	PLYISPIKL	IGVIAAVNMLT	VSSAVL G	AF F
LHR	VLI L	NI IMG MT	LFVLLT	RYK TVPRF	MCNLSFADFCM	LYLLL S	SQ K
FSHR	VLI F	SI ITG II	LVIIT	QYK TVPRF	MCNLAFADLCI	IYLLL S	IH K
TSHR	IVV FVSL	LLG VF	LLILLT	HYK NVPRF	MCNLAFADFCM	MYLLL S	LY H

	TM 3		□				
LGR4	GRFAEFG W	E S KV	SLA S	SA FL	LAAV	SVFAKDLMKHGKSSH	QF
LGR5	GSFARHGAW EN	V HVI	LSI	S FL	LAA	GFSVKYSAKFET	APPSSL
LHR	GQYYNHA D	Q S ST	FT	L YT	VIT	WHTITYAIHLDQ	LR HA
FSHR	SQYHNYA D	Q A DA	FT	L YT	AIT	WHTITHAMQLDC	VQ HA
TSHR	SEYYNHA D	Q P NT	FT	L YT	VIT	WYAITFAMRLDR	IR HA

**FIG. 6 (CONT)**

8/8

## TM 4

LGR4	QVAALLALLGAAVAGCF	FHGGQ SASPL	FPTGETPSLGFVTLVL	SL LLMA
LGR5	KVIIILLCALLALTM	AV L G K GASPL	LPFGEPMSTMG MVALIL	SLC LMMT
LHR	ILIMLGGWLFSSLI	ML V V N MKVSI	F MDVETTLSQV ILTILI	VV FIIC
FSHR	ASVMVMGWIWAFAA	LF IF I S MKVSI	MDIDSQL VMSLLV	VL VVIC
TSHR	CAIMVGGWVCCFLL	LL V I S AKVSI	MDTETPLALA IVFVLT	IV VIVC

## TM 5

LGR4	II T L CNL-EKEDLENSQSSVI	HV W	NCIFFC VA FSFAPLITAIS SPEI
LGR5	IA T L CNL-DKGDLENIW	CSMV HI L L	NCILNC VA LSF SLINLTF SPEV
LHR	AC I I FAVRNPELMATNK	TKIA KM I	DFTCMA IS FAI AAFKVPL TVTN
FSHR	GC IHI LTVRNPNIVSSSS	TRIA RM M	DFLCMA IS FAI ASLKVPV TVSK
TSHR	CCHV I ITVRNPQYNPGDK	TKIA RM V	DFICMA IS YAL AILNKPL TVSN

## TM 6

LGR4	M SVTLI F LPA L	V VF N
LGR5	I FI LVVV LPA L	L IL N
LHR	S VL VL Y INS A	F AI T
FSHR	A IL VL H INS A	F AI T
TSHR	S IL VL Y LNS A	F AI T

## C-terminal tail

LGR4	PK KE WKL KRRVTRKHGSVSVSISSQGGCGEQDFYYDCGMYSHLQGNLTVCDCCESFL
LGR5	PH KE LVS RKQTYVWTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSSITYDLPPSS
LHR	KT QR FFL LSKFGCCKRRAELYRRKDFSAYTSNCKNGFTGSNPKPSQSTLKLSTLHCQG
FSHR	KN RR FFI LSKCGCYEMQAQIYRTETSSTVHNTHPRNGHCSSAPRVTNGSTYILVPLS
TSHR	KA QR VFI LSKFGICKRQAQAYRGQRVPPKNSTDIQVQKVTHDMRQGLHNMEDVYELI
LGR4	LTKPVSKHLIKSHSCPVLTAASCQRPEAYWSDCGTQSAHSDYADEEDSFVSDSSDQVQA
LGR5	VPSPAYPVTECHLSSVAFVPC
LHR	TALLDKTRYTEC
FSHR	HLAQN
TSHR	ENSHLTPKKQGQISEEYMQTVL
LGR4	CGRACFYQSRGFPLVRYAYNLQRVRD

FIG. 6 (CONT)